Most Correct Answers: **#6**   Least Correct Answers: **#16**

1. **Complete link is**

1/12   (A)   the smallest distance between an element in one cluster and an element in the other

2/12   (B)   the largest distance between an element in one cluster and an element in the other

3/12   (C)   the average distance between an element in one cluster and an element in the other

5/12   (D)   distance between the centroids of two clusters

0/12   (E)   I do not know

2. **Which of these is the type of hierarchical clustering?**

3/12   (A)   Agglomerative Methods

3/12   (B)   Divisive Methods

2/12   (C)   Both

3/12   (D)   I do not know

3. **This function can be used to perform k-means clustering in R**

8/12   (A)   kmeans()

1/12   (B)   kclust()

2/12   (C)   kmenscl()

0/12   (D)   I do not know

4. **Do we need to worry about scaling in clustering with one single attribute?**

5/12   (A)   Yes

5/12   (B)   No

1/12   (C)   I do not know

5. In clustering, we seek to partition observations into distinct groups so that ...

7/12 **A** the observations within each are quite similar to each other, while observations in different groups are quite different from each other.

0/12 **B** the observations in different groups are quite similar to each other, while observations in the groups are quite different from each other.

1/12 **C** the observations in different groups and within the groups are quite similar to each other

1/12 **D** the observations in different groups and within the groups are different from each other

2/12 **E** I do not know

6. In k-means clustering

8/12 **A** we seek to partition the observations into a pre-specified number of clusters

2/12 **B** we do not know in advance how many clusters we want

1/12 **C** we end up with a tree-like visual representation of the observations

0/12 **D** I do not know

7. On each steps K-means algorithm

6/12 **A** is maximizes between-cluster variation and minimizes within-cluster variation

3/12 **B** is minimizes between-cluster variation and maximizes within-cluster variation

0/12 **C** depends on situation

2/12 **D** I do not know

8. The iteration in k-means stops

4/12 **A** until the cluster assignments stop changing

4/12 **B** until user-defined accuracy is achieved

0/12 **C** until some number of iterations

2/12 **D** all of the answers

1/12 **E** I do not know

9. In dendrogram

1/12 **A** observations that fuse at the very bottom of the tree are quite similar to each other, whereas observations that fuse close to the top of the tree will tend to be quite different

4/12 **B** observations that fuse at the very bottom of the tree are quite different, whereas observations that fuse close to the top of the tree will tend to be similar to each other

3/12 **C** observations that fuse at the very bottom of the tree and close to the top of the tree are quite similar to each other
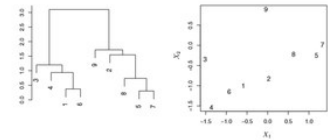
3/12 **D** I do not know

10. How many observations do we have in this agglomerative clustering?

6/12 **(A)** 9

1/12 (B) 20

0/12 (C) 90

1/12 (D) N

3/12 (E) I do not know
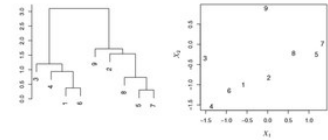
11. figure that observations 9 and 2 are quite similar to each other on the basis that they are located near each other on the dendrogram

7/12 (A) TRUE

3/12 **(B)** FALSE

1/12 (C) I do not know

12. We cannot draw conclusions about the similarity of two observations based on their proximity along the

4/12 **(A)** horizontal axis

0/12 (B) vertical axis

6/12 (C) both

1/12 (D) I do not know

13. One single dendrogram can be used to obtain any number of clusters

6/12 **(A)** True

3/12 (B) False

2/12 (C) I do not know

14. In agglomerative clustering, we start from

4/12 **(A)** the bottom of the dendrogram

6/12 (B) the top of the dendrogram

1/12 (C) I do not know

15. **When every object belongs to every cluster with a membership weight that is between 0 and 1, the clustering is called**

4/12 (A) Fuzzy

5/12 (B) Overlapped

2/12 (C) Exclusive

0/12 (D) I do not know

16. **Medoid**

0/12 (A) must be an actual data point

1/12 (B) almost never corresponds to an actual data point

1/12 (C) never corresponds to an actual data point

3/12 (D) can correspond to an actual data point

6/12 (E) I do not know

17. **K means perform badly on**

3/12 (A) Data with different sizes

0/12 (B) Data with different densities

1/12 (C) Data with non-spherical shapes

5/12 (D) All of the answers

1/12 (E) none

1/12 (F) I do not know

18. **The goodness of clustering structure can be checked**

4/12 (A) By adding new features

5/12 (B) By using different algorithms

1/12 (C) Both

0/12 (D) Cannot be checked

1/12 (E) I do not know

19. **K-means is sensitive to initial points**

7/12 (A) True

4/12 (B) False

0/12 (C) I do not know

20. We can perform the cluster analysis in R using the function

8/12 Ⓐ hclust

0/12 Ⓑ hierclust

3/12 Ⓒ clustering

0/12 Ⓓ I do not know